A template privacy protection scheme for fingerprint minutiae descriptors

Leila Mirmohamadsadeghi and Andrzej Drygajlo

School of Engineering, Swiss Federal Institute of Technology Lausanne (EPFL)

Lausanne, Switzerland

leila.mirmohamadsadeghi@epfl.ch, andrzej.drygajlo@epfl.ch

Abstract: It is important in biometric person recognition systems to protect personal data and privacy of users. This paper introduces a new mechanism to revoke and protect fingerprint minutiae information, which can be used in today's security-aware society. The recently developed minutiae cylinder code (MCC), which provides rotation and translation invariant descriptors for accurate fingerprint recognition by describing minutiae neighborhoods with respect to each other, is used as baseline fingerprint descriptor. A hybrid scheme combining a transformation and a user key is designed to provide the MCC-based fingerprint representation with revocability and irreversibility properties for template privacy protection across multiple applications. Furthermore, using the publicly available FVC datasets, it is demonstrated that the designed scheme improves the baseline accuracy of fingerprint recognition using the MCC method.

1 Introduction

Biometric template privacy protection is gaining importance with the widespread use of biometric person recognition and an increase in the awareness on related privacy issues. Biometric template privacy protection should ensure that an individual’s biometric characteristic is only available in the form of a template which is diverse and thus revocable if the template is compromised. Furthermore, the template must be irreversible to the original capture and must preserve the accuracy of the underlying recognition system [JNN08].

Recently, a specific scheme was designed to provide irreversibility for the minutiae cylinder code (MCC) representation, which involves the quantization of the Karhunen-Loeve transform [FMC12]. However, this scheme is not designed to provide diversity and revocability.

In the present paper, a novel method is proposed to provide template privacy protection for the minutiae cylinder code (MCC) representation [CFM10], which provides diversity, revocability and irreversibility properties, without degrading the baseline recognition accuracy. Fingerprint recognition using the MCC representation is chosen because it is alignment-free and computationally light.

The remainder of this paper is organized as follows. In Section 2, a brief description of the baseline fingerprint recognition system is provided. In Section 3, the proposed privacy protection scheme is presented. Experiments to assess performance of the protection
scheme in terms of accuracy are depicted in Section 4, while security aspects are addressed in Section 4.2. Conclusions are drawn in Section 5. Note: a complete version of this paper can be found at http://infoscience.epfl.ch/.

2 Fingerprint minutiae descriptors based on the minutiae cylinder code (MCC)

The minutiae cylinder code (MCC) [CFM10] is a recent fingerprint description method, which presents the advantages of both nearest-neighbor-based and fixed-radius-based minutiae description methods and is considered as the state-of-the-art in minutiae descriptor design [FZ11]. This method takes as input a set of standard ISO minutiae [ISO05], and creates for each minutia, a descriptor based on its distance to neighboring minutiae and their angular differences. This descriptor is of fixed length, robust to rotations and translations and skin distortions and is computed in a fast manner. The output consists of a template, which contains a descriptor for each minutia. This descriptor is a linearized cylinder whose discretized volume represents weighted spatial and angular distances of each minutiae to its neighbors. In order to compare two such templates within a recognition system, several comparison measures were originally introduced. In this paper, the local similarity sort (LSS) method is chosen among the others because it requires the least extra information about the original minutiae set when performing cylinder set comparison. The LSS comparison method computes all two by two distances of the cylinders and provides a similarity score based on the closest cylinder matches and angular distances of minutiae pairs. A slightly different version of the MCC method is used in this work. With respect to the original method, cylinder and cell validities are not considered and the weighting functions used to compute the spatial and angular contributions of neighboring minutiae are discretized.

3 A privacy protection scheme for the MCC templates

According to template privacy protection requirements, it is desirable to create from a raw biometric sample, several diverse and revocable templates which are irreversible to the original biometric characteristic and support accurate recognition. The baseline accuracy corresponds to recognition without template protection. It is thus required that template protection does not degrade this accuracy. Revocability is achieved when including a revocable component into the template, as original biometric characteristics are not revocable. Therefore, the privacy protection scheme presented in this paper is a hybrid two-factor scheme which transforms the template with a user key in a revocable and non-reversible manner. The key is assumed secret in the baseline protected system, but it is shown that even if the key is compromised or lost, the original biometric characteristic remains protected. The proposed solution takes root from cryptographic primitives such as the hard problem of square root modulo a composite number [HPS08]. During the modulo operation the quotient is dismissed, it is impossible to reconstruct accurately the original value. Also, as in filtering techniques, e.g., mean filtering, where the underlying operation replaces one value by a value inferred from several original values, the output is more ho-
heterogeneous with respect to the input and removes small variations such as noise. For this reason, the transformation presented here includes a step in which two original values are added to each other. The reduced intra-class variations are a positive step in working with biometric data, which suffer from inherent variations.

The output of the MCC descriptor creation is a template $T$ for every fingerprint, which contains a fixed length descriptor, referred to as cylinder $TC$, for each minutia. In order to dismantle the original structure of the cylinders, two-by-two elements are summed based on a user key, which is a permutation of the cylinder indexes. The remainder of the division of the square of this sum by a given parameter is then considered as a new revocable value as summarized in Equation 1. A binarization step is performed on the transformed values to introduce irreversibility and quantization, which is necessary in presence of the intrinsic intra-class variations of biometric data. Because of the small values of the baseline MCC template, a multiplication factor $A$ is used to adapt the size of the argument. A user key $k$, which is a random permutation of the cylinder indexes, is employed to specify the order of the summations of two-by-two elements. Changing this key ensures that it is possible to create several diverse instances of one biometric characteristic and allows to implement revocability through key management. The revocable template, which is the output of the method presented in this paper, is denoted by $RT$ and its cylinders as $RTC_s$.

The parameters of the protection operation include the multiplication factor $A$, the user key $k$, and the divisor $n$.

$$RTC[i] = B((A(TC[k(2i-1)]) + TC[k(2i)])^2 \ mod \ n)$$

for $i = 1, \ldots, nb_{elements}(TC)/2, \ \forall TC \in T,$ \hspace{1cm} (1)

where $B(v)$ binarizes each element of the descriptor by means of a threshold:

$$B(v) = \begin{cases} 1, & \text{if } v > t \\ 0, & \text{if } v \leq t. \end{cases}$$ \hspace{1cm} (2)

The values of $A$ and $n$ must be in accordance with each other in order for the modulo operation to be meaningful. It is observed that $\max(TC[k(i)] + TC[k(i+1)]) = 2$, because the output of the MCC method yields descriptors whose values are normalized between 0 and 1. In order for the modulo operation to yield meaningful results, $A$ and $n$ must be chosen such that $n < (2A)^2$. Furthermore, if the argument (i.e., $(2A)^2$) is very large compared to the divisor $n$, the discriminatory power of the biometric information is lost. Empirically, it is observed that the argument must not be more than three orders of magnitude larger than the divisor, i.e., $n < (2A)^2 < 10^3 n$. The threshold $t$ is determined based on the values of the transformed descriptors (which in turn depend on $A$ and $n$) and is empirically fine-tuned.

Given that the same transformation is applied on every cylinder in the template, it is possible to use the original LSS comparison measure of the MCC method [CFM10]. The LSS matching is based on computing the Euclidean distance of two cylinders. Given that by using the same key, the elements of the cylinders are shuffled in the same manner, their element by element distance does not vary. Due to the properties of Euclidean spaces and the nature of the transformation in Equation 1, a correlation exists between the distance of two descriptors before and after the transformation. This correlation is later
observed in the obtained results.

4 Experiments, performance results and discussion

In order to evaluate the proposed privacy protection scheme, the public and widely used FVC2002 [MMC+02] and 2004 [MMC+04] databases are used, which contain each, 8 impressions from 100 fingers. The minutiae of every fingerprint, formatted according to the ISO standard [ISO05], are extracted using the open source FingerJetFXOSE software by DigitalPersona [dig]. The original FVC protocol is used to generate genuine and impostor scores. The MCC algorithm is implemented according to Section 2 and is hereafter referred to as the “baseline MCC”. The MCC double-valued cylinder creation and matching parameters are $n_s = 16$, $n_d = 8$, $R = 75$, $\sigma_s = 6$, $\mu_\psi = 0.005$, $\sigma_D = 0.4363$, $min_{VC} = 20$, $min_M = 1$, $min_{ME} = 20\%$, $min_{np} = 3$, $max_{np} = 10$, $\mu_P = 10$, $\tau_P = 0.4$ and $\Delta \Theta = 2.35$. The size of each descriptor in the template using these parameters is 2048 elements. The minutiae extractor is modified to allow template creation for images with any number of minutiae.

4.1 Accuracy of the transformed templates

In order to assess accuracy changes after applying the template privacy protection method introduced in this paper, the genuine and impostor distributions are displayed in Figure 1 (a) for the two cases of recognition using the baseline MCC templates and recognition using the protected MCC templates on the FVC2002 DB1 in the case where each identity is assigned a different key. On the corresponding detection error trade-off (DET) curves in Figure 1 (b), it can be observed that recognition using protected MCC templates yields better overall separation and lower false accept rates (FAR) and false rejection rates (FRR) than recognition using the baseline MCC templates. Consequently, the equal error rate (EER) (the operating point at which FAR = FRR) is lowered as well. Corresponding error rates are reported in Table 1 for three operating points, as well as genuine/impostor class separation computed using the characteristics of Gaussian curves fitted to the actual distributions [MR05]. The EERs for the FVC2002 and 2004 databases (unseen data) are reported in Table 3. The parameters $A$ and $n$ (Equation 1) are set to $5 \times 10^3$ and $10^6$ respectively in order for the modulo operation to be meaningful while preserving the discriminatory power of the biometric information. Furthermore, in order to verify the validity of the conditions given for the values of $A$ and $n$ with respect to each other, several combinations were empirically chosen and tested. The threshold $t$ is set to $10^5$. However, this value is not optimal and other values within the range of $n$ are empirically tested in Table 2.

<table>
<thead>
<tr>
<th>Method</th>
<th>EER</th>
<th>FRR@FAR 1%</th>
<th>FRR@FAR 0.1%</th>
<th>Class separation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline MCC</td>
<td>1.21%</td>
<td>1.21%</td>
<td>2.25%</td>
<td>0.89</td>
</tr>
<tr>
<td>Protected MCC</td>
<td>0.72%</td>
<td>0.67%</td>
<td>1.39%</td>
<td>1.43</td>
</tr>
</tbody>
</table>

Table 1: Recognition results. FVC2002 DB1 images.
Figure 1: (a) Genuine and impostor distributions for the cases of recognition using the baseline MCC templates and the transformed MCC templates. FVC2002 DB1 images.; (b) DET curves for recognition using the baseline MCC templates and the transformed MCC templates. FVC2002 DB1 images.

<table>
<thead>
<tr>
<th>Threshold $t$</th>
<th>$10^5$</th>
<th>$2 \times 10^5$</th>
<th>$4 \times 10^5$</th>
<th>$6 \times 10^5$</th>
<th>$8 \times 10^5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>EER</td>
<td>0.72%</td>
<td>0.65%</td>
<td>0.49%</td>
<td>0.28%</td>
<td>0.36%</td>
</tr>
</tbody>
</table>

Table 2: Recognition results for different values of the binarization threshold $t$. FVC2002 DB1 images.

From the results presented in Figure 1, it can be concluded that the protection scheme presented in this paper improves the overall verification performance. This observation is made through the lowered error rates as well as an increased genuine/impostor class separation. This phenomenon is explained by the two-factor authentication. The user key provides extra discriminative information to the templates. It must be noted that experiments further in this paper (Section 4.2) show that by setting all keys to one unique and universal key, it is the biometric information being recognized and not the key. Another positive aspect of this privacy protection scheme is that the transformation reduces the size of the template by half. This property is beneficial when considering large databases as well as applications with reduced resources such as smart cards. Note: a comparison with the current state of the art can be found in the full version of this paper (http://infoscience.epfl.ch/).

4.2 Diversity and revocability of transformed templates

In targeted working conditions of the biometric recognition system, where the protection scheme is implemented and working in its normal mode, each user has his own key. This key is first used to enroll (enr) his template in the database and is later reused to reproduce the transformation during verification (ver). The operating point decision thresholds are determined in this scenario, which is referred to as same key enr/ver.

If a template is revoked, the corresponding key is black-listed and a new template is generated using a different key. The two templates must be different in order for the old template to be nullified. This scenario is referred to as different key enr/ver and corresponds to the case where a user needs several different instances of his biometric for use in different applications. Given that the original template consists of descriptors with $m = nb_{\text{elements}}(TC)$ elements ($m = 2048$ in this paper), $m!$ different keys can be gener-
ated. However, since the two elements of $TC$ specified by $k$ are added to each other before undergoing any other operations, the order of the elements does not matter. For example, if $nb_{elements}(TC) = 8$, then keys $k_1 = 37814256$ and $k_2 = 73814256$ yield the same protected template. They are equivalent to all other keys which have the same indices in the consecutive odd and even positions, regardless of the order of the indices within the pair. A “pair” here refers to an odd and even position in the key, for example $k[1]$ and $k[2]$ are a pair, as well as $k[3]$ and $k[4]$. If $k$ has $m$ elements, then there are $m^2$ possible permutations within the pairs since every pair has two permutations. The $2^{m/2}$ permutations yield equivalent keys as only the position within a pair varies from one key to another, not affecting the outcome of the transformation. These $2^{m/2}$ keys are a category of equivalent keys. Therefore in order to compute the number of distinct keys (which yield different protected templates through the transformation), the total number of permutations must be divided by the number of categories of equivalent keys, which results in $\frac{m!}{2^{m/2}}$ different keys. In order to ensure two diverse templates of one person are not similar, the pseudo-impostor accept rate (PIAR) is introduced to evaluate the proportion of successful pseudo-impostor attempts. A pseudo-impostor comparison is the comparison of two templates of one individual, which are generated using two distinct keys. It can be seen in Table 4 and Figure 2, that it is difficult to use a template generated with the wrong key for positive verification. If a user loses his key, it is known by an adversary. This scenario is simulated by using the same key for enrollment and verification, for all users, and is referred to as the stolen-token scenario. Furthermore, this scenario shows that it is the biometric information being verified, and not the key. These diversity and security testing scenarios are implemented according to [WH12].

<table>
<thead>
<tr>
<th>Threshold</th>
<th>same key ent/ver.</th>
<th>different key ent/ver.</th>
<th>stolen token</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FAR</td>
<td>FRR</td>
<td>FAR</td>
</tr>
<tr>
<td>0.4489</td>
<td>1%</td>
<td>0.67%</td>
<td>1.03%</td>
</tr>
<tr>
<td>0.4518</td>
<td>0.72%</td>
<td>0.72%</td>
<td>0.66%</td>
</tr>
<tr>
<td>0.4673</td>
<td>0.1%</td>
<td>1.39%</td>
<td>0.06%</td>
</tr>
<tr>
<td>0.4710</td>
<td>0.04%</td>
<td>1.6%</td>
<td>0.06%</td>
</tr>
<tr>
<td>0.4934</td>
<td>0</td>
<td>2.71%</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4: Error rates at different operating points under different diversity and security assumptions. FVC2002 DB1 images.

It can be concluded from the observations in Table 4, that it is possible to define an operating point for the privacy protection scheme presented in this paper, such that, a very limited number of impostors (and pseudo-impostors) gain access to the system with an unauthorized key. It must be noted that the security of this scheme is dependent on the decision threshold and offers flexibility in various operational scenarios.
4.3 Irreversibility of the transformed templates

In order to effectively protect a person’s biometric characteristic, it must not be possible to deduce said characteristic from the transformed template, even when all parameters, including the key, are known. In the transformation presented in this paper, non-reversibility is provided by the fact that during the transformation in Equation 1, part of the data is discarded through the modulo and quantization, which does not allow exact reconstruction even if all parameters are known. However, it is possible to create an approximation of a binary MCC template given a transformed template and the corresponding key as described in Equation 3:

\[
TC_R[k(2i)] = \begin{cases} 
1, & \text{if } RTC[i] = 1 \\
0, & \text{otherwise and}
\end{cases} \\
TC_R[k(2i + 1)] = \begin{cases} 
1, & \text{if } RTC[i] = 1 \\
0, & \text{otherwise.}
\end{cases}
\]

\[i = 1, \ldots, nb\text{elements}(RTC), \quad \text{and } \forall RTC \in RT. \quad (3)\]

\(TC_R\) denotes the approximated MCC templates and is compared to the baseline binary MCC template, for all users of a database in Figure 2 (b). It can be seen and concluded that these approximated templates do not resemble the baseline MCC templates and thus cannot be systematically used to extract useful minutiae information. In fact, when compared with genuine templates, the reconstructed templates are similar to impostor templates.

5 Conclusions and Future Work

In this paper, a novel template privacy protection technique for the MCC representation of fingerprint minutiae templates was presented. The proposed hybrid, two-factor technique combining a transformation and a user key, provides diversity, revocability, and irreversibility for the MCC descriptors with respect to the original minutiae information while improving the accuracy in recognition. Furthermore, the proposed technique reduces the template size by half.
Future work will include study and modification of the transformation in order to extend of the ideas presented in this paper to other modalities than fingerprints.

References


